Eur J Appl Physiol (2006) 96: 146–156
DOI 10.1007/s00421-004-1262-7

ORIGINAL ARTICLE

**Peter Schenk · Andreas Klipstein · Susanne Spillmann
Jesper Strøyer · Thomas Laubli**

The role of back muscle endurance, maximum force, balance and trunk rotation control regarding lifting capacity

Accepted: 11 October 2004 / Published online: 20 November 2004

Abstract Evaluation of lifting capacity is widely used as a reliable instrument in order to evaluate maximal and safe lifting capacity. This is of importance in regard to planning rehabilitation programs and determining working ability. The aim of this study was to investigate the influence of basic functions on the lifting capacity measured by the progressive isoinertial lifting evaluation (PILE) and the functional capacity evaluation (FCE) tests in a lower (floor to waist) and an upper (waist to shoulder) setting and compare the two test constructs. Seventy-four female subjects without acute low back pain underwent an examination of their lifting capacities and the following basic functions: (1) strength and endurance of trunk muscles, (2) cardiovascular endurance, (3) trunk mobility and (4) coordination ability. A linear regression model was used to predict lifting capacity by means of the above-mentioned basic functions, where the *F* statistics of the variables had to be significant at the 0.05 level to remain in the model. Maximal force in flexion showed significant influence on the lifting capacity in both the PILE and the FCE in the lower, as well as in the upper, lifting task. Furthermore, there was a significant influence of cardiovascular endurance on the lower PILE and also of endurance in trunk flexion on the lower FCE. Additional inclusion of individual factors (age, height, weight, body mass index) into the regression model showed a highly significant association between body height and all lifting tasks. The r^2 of the original model used was 0.19/0.18 in the lower/upper FCE and 0.35/0.26 in the lower/upper PILE. The model r^2 increased after inclusion of these individual factors to between 0.3 and 0.4. The fact that only a limited part of the variance in the lifting capacities can be explained by the basic functions analyzed in this study confirms the assumption that factors not related to the basic functions studied, such as lifting technique and motor control, may have a strong influence on lifting capacity. These results give evidence to suggest the inclusion of an evaluation of lifting capacity in clinical practice. Furthermore, they raise questions about the predictive value of strength and endurance tests in regard to lifting capacity and work ability.

Keywords Body height · Endurance · Gross motor coordination · Lifting capacity · Trunk strength

P. Schenk (✉) · T. Laubli
Institute of Hygiene and Applied Physiology,
Swiss Federal Institute of Technology Zurich,
8092 Zurich, Switzerland
E-mail: schenk@iha.bepr.ethz.ch
Tel.: +41-1-6324809
Fax: +41-1-6321173

A. Klipstein · S. Spillmann
University Hospital Zurich, Department of Rheumatology
and Institute of Physical Medicine, Zurich, Switzerland

J. Strøyer
Department of Physiology,
National Institute of Occupational Health,
Lersø Parkalle 105, 2100 Copenhagen, Denmark

Introduction

Evaluation of maximal lifting capacity is commonly used in order to determine the work-related physical capacity of workers in physically demanding jobs, to plan rehabilitation programs and to decide on the return to work or payment of compensation. Knowledge of work-related physical capacity is very important to determine safe, tolerable levels of load for the injured worker with respect to work demands, to predict when an individual is ready to return to work after injury, and to evaluate the effectiveness of rehabilitation programs.

Originally, the focus of functional capacity evaluation was on the assessment of vocational functions, especially in jobs requiring large components of physical load. Nevertheless, the value of functional assessments in non-working populations has been established (Jones and Kumar 2003). The high reliability of these lifting tests has been shown by Gross and Battie (2002) and Mayer et al. (1988a). Since lifting tests consist of simulated manual material handling tasks, high validity could be expected, but scientific evidence for their validity is unknown.

As reviewed by Mital et al. (1993), lifting capacity is limited by static and dynamic muscle strength and the resulting compression of the lumbar discs is considered as a main risk factor for injuries. Based on the strength of the trunk muscles and acceptable compression forces on the lumbar discs, limits of safe lifting have been calculated using biomechanical models.

Nowadays, there are two different approaches to experimentally assess maximum lifting capacity, namely psychophysical and kinesiophysical tests. Both consist of repeatedly lifting boxes of increasing weight. In the functional capacity evaluation (FCE) by Isernhagen Work Systems, which is a kinesiophysical test, the administering therapist is in control and tasks are stopped when biomechanical, respiratory or cardiovascular signals of maximal effort are observed or when the safe lifting limit is reached (Isernhagen 1991, 1992). If the subject wishes to stop the test before reaching the functional limit, the subject is motivated by the therapist to continue with the test. Therefore, by means of the FCE, the administering therapist can identify what the subject is able to do, and what the subject is willing to do.

The progressive isoinertial lifting evaluation (PILE) described in Mayer et al. (1988a) is a psychophysical test. The development of psychophysical lifting tests is summarized in Troup et al. (1987). The PILE is stopped when time limits for four lifting movements are exceeded or when the cardiovascular limit is reached. The test can also be stopped by the subject when he or she is fatigued or feels pain. It measures the subjects' ability to cope with a physical demand and the maximum weight lifted does not necessarily correspond to an acceptable and save workload.

In order to optimize rehabilitation programs and exercise therapy with regard to return to work, it is essential to know whether the lifting capacity can be explained and predicted by distinct basic functions of the trunk. Otherwise, lifting capacity would be determined by individual traits and characteristics (Mayer et al. 1988b). The aim of our study was to assess the predictive value of basic functions thought to be relevant for lifting capacity measured with the FCE and PILE tests in a lower (floor to waist) and an upper (waist to shoulder) setting: general fitness, strength and endurance of the muscles involved in moment production in the sagittal plane, mobility, as well as gross motor control and coordination ability. Furthermore, we wanted to compare the two different psychophysical and kinesiophysical test constructs.

The data analyzed and described in this paper were collected during a set of examinations within the European cost-shared project NEW (neuromuscular assessment in the elderly worker). The psychosocial factors also assessed in the study are not included in this analysis and will be published elsewhere.

Methods

Subjects

A total of 114 female volunteers participated after giving their informed consent. The subjects participating in this study were recruited from the staff of the University Hospital, Zurich. The inclusion criteria were: women in nursing or administrative occupations aged from 45 to 62 years who worked at least 20 h/week without any interruptions of more than 1 month during the previous 5 years. Subjects who were pregnant, used prescription heart or lung medications, or with hypertension, angina pectoris, fever, or a resting heart frequency of more then 120 beats per minute were excluded from the tests. For this analysis, subjects that suffered from lower back pain (LBP) for more than 30 days during the last year, as well as subjects who reported current LBP stronger than 4 on the numeric rating scale of 1–10 during the previous 7 days, were excluded.

Finally, 74 subjects with a mean age of 52.3 (4.8) years, body mass of 65.2 (10.2) kg and height of 164.6 (6) cm were included in this analysis. Forty-one of the subjects had not experienced LBP within the previous 12 months, 16 subjects reported pain over a period of 1–7 days and 17 subjects reported 8–30 days with LBP within the previous 12 months. Twenty subjects experienced low back pain during the 7 days prior to the first experiment. The average pain level was 1.4 (1.8) on the numeric rating scale. All participants gave their written informed consent and the study was approved by the ethical committee.

Study protocol

The whole testing consisted of two sessions of approximately 2.5 h each. A warm-up consisting of lateral and frontal bending, as well as trunk rotation movements, was supervised by the physiotherapist and performed before the functional tests. The first session started with a clinical examination by a physician. Subsequently, lumbar flexibility, balance and aerobic capacity were tested. This was followed by a questionnaire to be completed by the subject concerning her physical and psychosocial work demands, psychosocial factors and pain in different regions. The first session was completed by two elements of the FCE test according to Susan Isernhagen (Isernhagen Work Systems). The second session, usually performed the following week, started with a lumbar mobility analysis and a gross motor tracking test involving trunk rotation movements. Then, a stress and an eye-hand coordination test were performed on a computer. The endurance and maximal voluntary contraction (MVC) of the trunk extensor and flexor muscles and the MVC of the trapezius muscle were measured. The second session closed with the PILE test. Before starting the lifting tests, subjects were asked about fatigue or pain due to the previous tests, and, if required, a break was taken until full relaxation.

Two instruments were used to measure lifting capacity. During the FCE according to Isernhagen et al. (1999), the safe lifting capacity, i.e. the workload that can be handled without overexertion was assessed. The test consisted of an upper lifting test (waist to shoulder) and a lower lifting test (floor to waist). The height of the lower shelf was individually adjusted to achieve a horizontal position at the forearm and the upper shelf was adjusted to bring the handgrips of the box to eye height. The weight to be handled began at 2.5 kg and was increased progressively by 2.5 kg for each step. Each weight had to be lifted 5 times (waist to floor to waist) within 90 s. After each test sequence, a short break was taken to note the observations as well as the subjects' feelings of discomfort. The test was stopped when the observer detected unsafe technique or when the aerobic end point of 85% of the age-related maximal heart rate was reached. Maximal heart rate was calculated as 220 − age. Thus, the outcomes of the functional capacity evaluations were the maximal weight lifted safely, and the reason for the limitations. Test-retest reliability in LBP patients has been shown to be good with intraclass correlation coefficients (ICC) of 0.78–0.83 for lower and 0.81–0.84 for upper FCE (Brouwer et al. 2003; Gross and Battie 2002).

The PILE was used as a second test to quantify the subjects' physical lifting capacity and ability to cope with a physical workload. The PILE test was performed in a lower (floor to waist) and in an upper (waist to shoulder) setting. The shelf level for the waist and shoulder were set at 75 and 137 cm according to the test procedure of Mayer et al. (1988a). The starting weight of the box was 4 kg and was increased stepwise by 2.5 kg at each increment. The subjects were required to perform 4 lifting movements within 20 s for each weight. The test was stopped when the subject terminated it due to discomfort or by exceeding the time limit, when the aerobic endpoint of 85% of maximal heart rate (190 − age) was reached, or when the safe end point of 55–60% of the subject's body weight was reached. Therefore, the result of the PILE consisted of the maximum weight lifted, the final heart rate being the reason for stopping. Since functional and ergonomic aspects such as safe and controlled techniques were not taken into account during the PILE test, the results consist of values that can be compared with normative data, but no comparison can be drawn to real working capacity (Olivieri 1999). The test-retest reliability of the maximum weight lifted during the PILE test has been shown to be good with correlation coefficients of 0.87 for the lower (lumbar) and 0.93 for the upper (cervical) settings (Mayer et al. 1988a).

Factors presumed to determine the outcome of the lifting capacity

We assessed strength, endurance and mobility of the trunk, cardiovascular endurance and gross motor control as factors we expected to contribute to the lifting capacity. Several methods to the measure functional properties of the trunk have been published; we based our choice for the strength, endurance and mobility tests on the review of functional tests by Essendrop et al. (2002).

Endurance and maximum force of trunk muscles

The trunk extensor and flexor muscles are involved in the movements of lifting, as well as in the stabilization of the trunk during lifting tasks. Therefore the endurance and MVC of the trunk muscles were presumed to contribute a large percentage of the lifting capacity.

Endurance of trunk extensor muscles

The endurance of the trunk extensor muscles was assessed in a modified Sorensen test adapted from Ito et al. (1996). Subjects were placed in the prone position on a board of 70 cm in length at an ascending slope of 10° with the navel precisely over the edge of the sloping board. The subjects were asked to fold the arms across the chest and to lift the upper body to a horizontal position while the helper gently held the feet to the ground. In case of a deviation from the horizontal position the subject was asked to adjust her position. The test was stopped as soon as the upper body fell considerably lower than the horizontal position more than twice. The total time (up to 360 s) the subject could maintain the horizontal position was measured. The interobserver reliability of the modified Soerensen test (expressed as the Pearson correlation) has been shown to be 0.93 (Westhoff 1994).

Endurance of trunk flexor muscles

Evaluation of the static endurance of abdominal muscles was conducted according to Ito et al. (1996). Subjects were asked to lie in a supine position and to place the lower leg on a box, thus forming a 90° angle with the thighs. They were then required to fold the arms in front of the chest and to curl upwards just a few centimeters such that the scapulae were no longer in contact with the floor. The subjects were instructed to maintain this position with the upper back unsupported as long as possible, and they were encouraged to maintain the position. The test was stopped as soon as the scapulae touched the floor for the third time, or as soon as a 120 s maintenance time was reached. The duration that the correct position was maintained was measured. The intraobserver reliability (Pearson correlation coefficient) of the abdominal endurance test has been shown to be 0.93 (Hyytiainen et al. 1991).

Maximum contraction of trunk extensor and flexor muscles

Subjects with blood pressure higher than 100/160, acute signs of nerve root compression or diagnosed osteoporosis were excluded from the MVC measurements. To measure the maximum voluntary isometric strength of the trunk extensor and flexor muscles, a frame was used where subjects were fixed in standing position with a strap around the hips at pelvis height. The subject stood with her face towards the equipment and the load cell was adjusted to the height of the subject's breastbone. The test procedure was adapted from Biering-Sorensen (1984). The subject was instructed to build up the force for 5 s, keep it for 2 s and then lower the force to zero. The contraction was repeated at least 3 times with breaks of 30 s between contractions. For all MVC measurements, the examiner encouraged the subjects to reach their maximum force. The maximum force levels (in Newton) were read from the force transducer (Mecmesin AFG-R 1,000 N). If the third MVC was 5% higher than the second, then a fourth measurement was done. If there was still an increase of at least 5% from the third to the fourth contraction, then the contraction was repeated for a last time and the highest value was taken as the result.

For measurement of the MVC in extension, the same procedure was applied as for flexion, with the subject standing with her back towards the equipment. The height of the load cells was adjusted to the subject's scapulae and the width of the load cells was adjusted so that they were in-between the margo medialis of the scapulae and the spine. The reliability of this experimental setup was not assessed, but Essendrop et al. (2001) have shown excellent reliability (ICC > 0.9) of strength measurements in trunk flexion and extension in a similar experiment with subjects pulling against a strap around the shoulders.

Mobility of the trunk

Adequate mobility of the trunk in both the lateral and in frontal directions was presumed to be an indispensable requirement for successful lifting. The mobility in the lateral and in the frontal direction was measured. A box to which a ruler with a slider was attached was used for both sets of measurements.

Measurement of frontal bending

Subjects stood on the box without shoes, with the tip of the feet touching the edge of the box, one foot placed on each side of the measurement ruler. The subjects were asked to bend slowly forward, as far as possible, pressing the ruler slide as far as possible, while keeping the knees extended. The difference from the fingertip to the floor was measured, where negative values indicate that the subject could reach further down than the tip of the toes during bending forward. Gauvin et al. (1990) found excellent reliability of fingertip to floor measurements, with an ICC of 0.98. Measurement of frontal bending was performed twice, intermitted by the first measurement of lateral mobility. The consistency of the two frontal mobility measurements, measured by the Cronbach alpha coefficient, was 0.93.

Measurement of lateral bending

The mobility of lateral bending to the left and to the right was measured using the same box as for the frontal mobility assessment. The subject stood on the box, without shoes, with the side towards the measurement ruler and the palm turned towards the body, so that the middle finger was in a line with the ruler. The position of the tip of the middle finger was measured for the upright starting position and for the position in maximal lateral flexion.

Measurement of maximum bending to both sides was performed twice intermitted by the second frontal measurement. The range of motion in lateral bending was calculated as the difference from the starting position to the maximum lateral position. This test has been evaluated for reproducibility by Suni et al. (1996). The coefficient of variation for test-retest was 4.7% and inter-rater reliability was (ICC) 0.92. The two measurements to each side were aggregated to create one index for lateral mobility with a Cronbach alpha coefficient of 0.92, confirming the consistency of the scale.

Aerobic capacity

The tasks during FCE and PILE consisted in handling loads of increasing weight; it was presumed that the subject's weight-related maximal oxygen uptake (WR $\dot{V}O_{2max}$) reflects general fitness that contributes to the maximal weight that can be lifted. The maximal weight-related oxygen uptake was extrapolated from a submaximal cycle ergometer test according to Astrand (1970). The heart rate was measured by a Polar heart rate monitor and was checked every 15 s if it remained within the target range. Age-corrected and weight-related maximal oxygen uptake $\dot{V}O_{2age}$ [ml O_2/min/kg] was calculated according the formula

$$VO_2age = VO_2 * (-0.3637 * 1n(\text{Age}) + 2.1683)$$

derived from experimental data by Astrand (1970) where $\dot{V}O_2$ was the value estimated by the nomogram. The r^2 of this approximation was 0.9991. According to Gore et al. (1999), the standard error of submaximal determination of $\dot{V}O_{2max}$ is 0.85.

Gross motor control

Handling of weights requires a delicate coordination of the muscles involved in moment generation, stabilization of the trunk, and in compensation of the disturbance of the body's centre of gravity by the weight being handled. Gross motor control was assessed twofold: the SD of the centre of gravity in the sagittal plane as well as in the coronal plane was measured while standing on a force platform (Kistler type 9286), standing on both legs with the eyes open, standing on the dominant leg with the eyes open, and standing on both legs with the eyes closed. Each test was performed for 60 s and the standard deviation of the centre of gravity was calculated from the middle 40 s. High values correspond to a weak balance stabilization. The values corresponding to the three test sequences were summed to a global measure for balance. The consistency of this scale for balance, measured by the Cronbach alpha coefficient, was 0.78. The reliability of balance measurements made while standing on both legs was reported by Brouwer et al. (1998) to be poor, with an ICC of 0.45 with the eyes open and 0.38 with the eyes closed.

The gross motor coordination ability was assessed in a trunk rotation-tracking test. Subjects were fixed to a chair with a laser pointer attached to their chests. They were asked to track a bar moving in a horizontal plane back and forth at five different velocities. At each velocity, 10 movement cycles had to be performed. The movement frequency started at 0.0781 Hz and was doubled doubled until 1.25 Hz, adding up to 5 test sequences of 10 movement cycles each. The moving bar was adjusted to the expected amplitude of the trunk rotation movement, 45° to each side. Movements of C7 and the shoulders were recorded by a sonic ZEBRIS CMS20S device at a sampling rate of 50 Hz. The movements were decomposed into rotational movement, lateral bending and forward bending. The first and the last of the 10 movement cycles were excluded from the calculations. For each of the five movement velocities, the mean amplitude of the eight rotation movements to the right and to the left, as well as the mean peak angular velocities of the rotation movements to the right and to the left were calculated. These values were summed to a global measure for gross motor coordination ability. The consistency of this scale for coordination ability, as measured by the Cronbach alpha coefficient, was 0.86.

Statistics

For data analysis, the regression procedure in the SAS 8.02 for Windows package was used. Missing values from the single tests were replaced by the mean value of the corresponding variable. A linear regression model with maximum weight-related oxygen uptake (WR $\dot{V}O_{2max}$), endurance and maximum voluntary contraction of trunk muscles, flexibility in lateral and frontal bending and two measures of coordination ability as independent variables was used to determine the maximal lifting capacity. Due to deviations from a normal distribution, the logarithm of the endurance time for trunk flexion and extension was used in the model. The predictor variables were independent, with mutual correlations of <0.4 except for the endurance time in extension and in flexion, where the correlation was 0.56.

Lifting Capacity = WR $\dot{V}O_{2max}$ + Log(Endurance$_{Extension}$) + Log(Endurance$_{Flexion}$) + MVC$_{Extension}$ + MVC$_{Flexion}$ + Mobility$_{Frontal}$ + Mobility$_{Lateral}$ + Balance + Coordination

Table 1 Outcomes of the four lifting tasks (maximum weight lifted). *FCE* Functional capacity evaluation, *PILE* progressive isoinertial lifting evaluation

Test	*N*	Mean (SD) weight (kg)
Lower FCE	74	14.66 (2.99)
Upper FCE	74	9.29 (2.05)
Lower PILE	73	17.25 (4.89)
Upper PILE	72	11.36 (2.99)

The model was tested with the outcomes of the maximal lifting capacity measurements assessed in the two tests (PILE and FCE) in the lower and upper setting. Due to the limited number of subjects ($n=74$) compared to the extensive set of nine predictors in the original model, a stepwise model selection method was applied where F statistics for a variable to be added to the model were required to be significant at the 0.1 level and the F statistics of the variables remaining in the model were required to be significant at the 0.05 level. For all variables remaining in the model, estimated model parameters and standard errors as well as partial r^2 values were calculated.

Results

The maximum weight lifted, evaluated either by the FCE or the PILE, in both the lower and the upper settings, differed significantly. It ranged from 9.29 kg in the upper FCE to 17.25 kg in the lower PILE. The maximum weight lifted was significantly greater in the PILE than in the FCE and also significantly greater in the lower setting than in the upper setting, as tested by both the PILE and the FCE. The mean difference between the two lifting procedures was 2.6 kg in the lower and 2.1 kg in the upper setting. The outcomes of the lifting tasks are summarized in Table 1. The relationship between the two protocols and the two settings are provided in Table 2. The correlation of the four lifting tasks (PILE and FCE in both the upper and lower settings) with each other was moderate (0.3–0.5) with the lowest values observed for the comparisons of different height levels or different protocols.

The outcomes of the functional variables used in the regression model and the numbers of samples available are listed in Table 3. Differences in the number of

Table 2 Pearson correlation coefficients between the two lifting test protocols (FCE and PILE) in the lower and the upper setting in the study sample of 74 subjects

Test	Lower FCE	Upper FCE	Lower PILE	Upper PILE
Lower FCE	1	–	–	–
Upper FCE	0.53***	1	–	–
Lower PILE	0.43***	0.34**	1	–
Upper PILE	0.55***	0.41***	0.46***	1

*$P<0.05$, **$P<0.01$, ***$P<0.001$

Table 3 Number of samples, mean value and standard deviation (SD) of the variables included in the model to predict lifting capacity

Test	N	Mean (SD)
WR $\dot{V}O_{2max}$ (ml O_2/min/kg)	68	31.5 (7.6)
Endurance flexion time (s)	72	89.5 (58.9)
Endurance extension time (s)	68	157.9 (92.7)
MVC flexion (N)	74	258.9 (72.5)
MVC extension (N)	74	290.6 (128.1)
Lateral mobility (cm), the sum of two measurements for each side	73	65.9 (11.2)
Fingertip to floor distance (cm), the sum of two measurements	58	0.8 (16.3)
Balance index	59	6.7 (3.8)
Coordination index	52	1,213.5 (377.2)

available measurements are due to missing or invalid data, intermittent failure of some devices or exclusion of subjects from single tests due to medical conditions.

The results from the linear regression model for the four lifting capacity measurements are summarized in Table 4. The estimated parameter and its standard error are provided for each variable that remained in the model. Additionally, the r^2 value for the total model and the single predictor r^2 values and their significance levels are listed. There was a highly significant association between the MVC in trunk flexion with both the PILE and the FCE in the lower and the upper settings. Additionally, we found highly a significant influence of endurance (trunk extension) on the lower FCE and of aerobic capacity on the lower PILE. Furthermore, there was a significant negative influence of postural sway measured on the force platform on lifting capacity in the lower PILE, as well as an influence of endurance in trunk extension and lateral mobility on the upper PILE. Gross motor coordination ability as assessed in the trunk rotation test, mobility in the frontal plane and MVC in extension showed no influence on the lifting

Table 4 Estimated parameters and standard errors (in *parentheses*) of the linear regression model for upper/lower FCE and PILE. *EndExt* Endurance of trunk extensor muscles, *MaxFlex* maximal voluntary contraction of the trunk flexor muscles, $WR\dot{V}O_{2max}$ weight-related maximal oxygen uptake, *MobLat* mobility in lateral bending. The partial r^2 of the variables and their significance level are also given

	Lower FCE	Upper FCE	Lower PILE	Upper PILE
	0.18***	0.19***	0.35***	0.26***
EndExt				
Parameter estimate (SE)	1.69 (0.558)[a]			1.47 (0.527)
Partial R^2	0.11**			0.08*
MaxFlex				
Parameter estimate (SE)	0.01 (0.004)	0.01 (0.003)[a]	0.03 (0.006)[a]	0.01 (0.004)[a]
Partial R^2	0.07*	0.19***	0.18***	0.13**
Balance				
Parameter estimate (SE)			-0.28 (0.139)	
Partial R^2			0.03*	
WRVo2max				
Parameter estimate (SE)			0.24 (0.064)[a]	
Partial R^2			0.13***	
MobLat				
Parameter estimate (SE)				0.06 (0.027)
Partial R^2				0.06*

*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$
[a]Significant predictors for lifting capacity

Table 5 Pearson correlation coefficients of the maximal lifting capacity with individual factors potentially influencing the outcome

Test	BMI (kg/m^2)	Body mass (kg)	Height (cm)	Age (years)
Lower FCE	0.01	0.19	0.34**	−0.21
Upper FCE	0.01	0.19	0.41***	−0.18
Lower PILE	−0.18	0.04	0.34**	−0.26*
Upper PILE	−0.01	0.15	0.39***	−0.27*

* $P<0.05$, ** $P<0.01$, *** $P<0.001$

capacity in any of the test settings. The variables included predicted 18–35% of the total variance with the strongest relationships for lower and upper PILE.

Body mass and body mass index (BMI) did not correlate with the outcomes of the lifting tasks. Height correlated strongly with the outcomes of all four lifting tests. Age correlated negatively with the maximal lifting capacity. This correlation was statistically significant for the PILE in both the upper and the lower settings (see Table 5). It is of interest to determine if the relationships between the lifting capacity and the functional variables described here are influenced by the individual factors. Table 6 lists the estimated parameters, the standard errors and the partial r^2 values of the same model as described above, after inclusion of age, height, body mass and BMI. Parameters that were highly significant ($P<0.01$) in the original model changed only marginally.

Table 6 Estimated parameter, standard error and partial r^2 values of the regression model after the inclusion of individual factors. Height was a significant factor for all lifting tasks. There were marginal changes in the estimated parameters belonging to the functional variables (compare Table 4) indicating the independence of the functional variables from individual factors such as height

	Lower FCE 0.32***	Upper FCE 0.40***	Lower PILE 0.41***	Upper PILE 0.31***
EndExt Parameter estimate (SE) Partial R^2	1.83 (0.601)[a] 0.09**			1.29 (0.51) 0.06*
MaxFlex Parameter estimate (SE) Partial R^2		0.01 (0.002)[a] 0.19***	0.03 (0.006)[a] 0.18***	0.01 (0.004)[a] 0.08**
Balance Parameter estimate (SE) Partial R^2	-0.22 (0.106) 0.04*	-0.21 (0.066)[a] 0.06**	-0.39 (0.139)[a] 0.04**	
WRVo2max Parameter estimate (SE) Partial R^2			0.19 (0.063)[a] 0.13**	
MobLat				
Age				
Body height	0.24 (0.057)[a] 0.12***	0.14 (0.034)[a] 0.12***	0.23 (0.082)[a] 0.07**	0.16 (0.050)[a] 0.16**
Body weight		0.04 (0.021) 0.03*		
BMI [kgm^{-2}]	0.28 (0.092)[a] 0.06**			

* $P<0.05$, ** $P<0.01$, *** $P<0.001$

[a]Significant predictors for lifting capacity

Height was a strong predictor for lifting capacity in all lifting tasks.

Discussion

This study focused on the importance of distinct functional capabilities in predicting lifting capacity as determined by two commonly used test procedures which measure a subject's ability to cope with a workload (PILE) or the maximum safe weight levels for rare lifting (FCE). The main results were that the average maximum weight lifted in the PILE and the FCE were similar, although statistically different. There are no published experimental data concerning lifting capacity in a group comparable to ours, but the maximum weight lifted in our study sample corresponds to published and recommended limits of weight for rare lifting (Mital et al. 1993; Steinberg and Windberg 1994). Although both lifting protocols indicated similar values for lifting capacity, their correlation with each other was moderate. The other tests selected to describe trunk function only moderately predicted the outcome of the lifting tasks. This indicates that low back function cannot be described by single measures.

We compared the outcomes of the functional tests in our study group with previously published data for healthy subjects to make sure that the results from the regression model are not limited to a biased study population. The endurance time for isometric trunk extension in our tests was slightly longer ($P < 0.05$) than in data published by Ito et al. (1996), but there was no significant difference in trunk flexion endurance when comparing our data with the study by Ito et al. (1996).

There was no significant difference between our data for mobility in frontal bending (fingertip to floor distance) and those published by Kippers and Parker (1987) but the mobility in lateral bending for our group was significantly smaller than data published by Suni et al. (1996). Comparison of aerobic capacity with other samples is difficult due to its strong dependence on age and gender. Our values were smaller than those reported by Keller et al. (2001), who used the same test protocol but also included younger subjects as well as men. There are no normative data available for the balance index, coordination index and isometric contraction of the trunk muscles as assessed in our setting.

The sample group we tested was limited to employed women aged between 45 and 62 years, and current as well as severe low back pain patients were excluded from this analysis, resulting in a homogeneous group. A selection bias may be possible for the test group, since all participants were volunteers and the tests were run in their leisure time. Clearly, our study group is not based on a random sample, however, it includes both subjects with physically demanding (nursing) and sedentary (administrative) work. Thus we believe that generalization to women in the same age group (45–62 years) is justified.

All tests used in this analysis proved to be practical; they did not cause pain and there was a negligible risk of injury. Although the test sessions were long, subjects did not make use of additional breaks to recover from the single tests. Subjects with acute pain (higher than 4 on the numeric rating scale of 1–10) were excluded from this analysis, although in medically stable conditions pain is not a contraindication for the FCE (Hart et al. 1993). No subjects with medical conditions requiring special treatment were included in the test group. A limitation with regard to the development of the regression model consisted in the rather rough gradation of the outcome of the lifting tasks (in all tests, the weight was increased stepwise by 2.5 kg). Reducing the gradation is not practical because the test time would be dramatically increased and thus there would be a stronger influence of general fatigue. We also presume that in the FCE, the accuracy of estimating safety, partial overload and local fatigue by the observer is in the range of 2.5 kg.

It became clear that a set of many trunk functions tests could not predict the lifting capacity. MVC in trunk flexion was the only factor that had a significant influence on both lifting tests in both settings. The flexor muscles may act as a stabilizer in lifting tasks. Cholewicki et al. (1999) have investigated the role of intra-abdominal pressure (IAP) and of coactivation of abdominal muscles on lumbar stability in a simulation model; they concluded that both IAP and coactivation of abdominal muscles contribute to the stabilization of the lumbar spine. IAP could be of significant importance in tasks requiring the production of trunk extensor moments, where coactivation of abdominal muscles would counteract with the intended moment production. The contribution of the trunk flexor muscles to lifting tasks was also shown in an electromyography study by Gorelick et al. (2003). Since the contraction levels of the abdominal muscles during lifting tasks are rather low (de Looze et al. 1999), we concluded that the MVC in trunk flexion might partly reflect inter- and intramuscular coordination of the abdominal muscles leading to improved lifting capacity. Furthermore, in the setting we used to measure the maximal strength in trunk flexion (pressing with the breast bone against a force transducer) a high level of trunk coordination and stabilization was required. Therefore, it might not be maximum strength that determines lifting capacity, but the coordinative skills that are needed to achieve high force levels with the maximum force measurement method that was used in our study. In contrast to the MVC, endurance in trunk flexion showed no influence on the lifting tests. On the other hand, the endurance in extension showed some significant relationships with the lifting capacity.

There was a significant influence of aerobic capacity on the lifting capacity in the lower PILE. This outcome was expected since the PILE places more emphasis on general endurance than the FCE. In the FCE, the influence of aerobic capacity is limited since a period of 90 s can be used to accomplish 5 lifting cycles compared to 20 s for 4 lifting cycles in the PILE. This influence of aerobic capacity on the PILE could be overcome by means of breaks between the test sequences, which on the other hand, would lengthen the test procedures leading to adverse effects such as increased costs.

From a biomechanical point of view, the significant relationship of the lateral mobility with the lifting capacity in the lower PILE is not clear. We hypothesized that lateral mobility might partly reflect general mobility. Interpretation of the influence of balance on lifting capacity is difficult and of limited significance since the reliability of measurements of postural sway on the platform show a low reliability of 0.38–0.45 (Brouwer et al. 1998). A significant relationship between balance and lifting capacity was found in the lower PILE, which requires repeated rotational trunk movements.

The coordination index was found to have no influence on the lifting capacity. This could be due to the fact that our rotation test measured the agility and the coordination ability of the muscles contributing to trunk rotation, which cannot directly be transposed to the demands of lifting tasks, although at least the lower lifting tests required a trunk rotation under load in order to place the box beside the shelf. Furthermore, it is not clear whether or not the coordination pattern of the muscle groups involved in lifting movements can be observed by means of whole body motion analysis, or whether the motion of the single muscle groups should be observed. The use of ultrasound imaging methods (Bunce et al. 2004) or fine wire intramuscular EMG to assess activation patterns of single muscle groups could provide further insight into the role of muscular coordination on lifting capacity.

There was a significant ($P < 0.05$) negative correlation of age with lifting capacity measured by the PILE but not by the FCE. However, when considering the individual factors in the original model, age was not a significant factor for lifting capacity (see Table 5). It is interesting to note that the FCE, which reflects maximal safe lifting capacity and which was expected to be valid in respect to real working conditions, did not show an association with age.

There was significant influence of height on lifting capacity in both the PILE and the FEC and in the upper, as well as in the lower, settings. Other individual factors such as body mass and body mass index (BMI) showed no significant influence on the lifting capacity. We expected an influence of height mainly on the upper PILE since in the PILE protocol the shelves were adjusted to a height of 137 cm, whereas in the FCE the shelves were adjusted to the individual's height. Therefore, in the PILE smaller persons had to lift the box relatively higher than tall subjects. The height of the shelves in the FCE was 132.5 (6.9) cm and the height of the subjects was 164 (4) cm.

We expanded the original model by adding individual factors such as age, body mass, height and BMI. Height became a highly significant predictor for lifting capacity in all tasks (see Table 6). There were only marginal changes in the factors that were highly significant in the original model (compare Table 4). This shows that height was not a covariable for the basic functions measured, but a separate factor influencing the lifting capacity. There is a lack of knowledge regarding the importance of height in regard to trunk functions. A single study (Luk et al. 2003) has shown a weak correlation between height with lifting capacity. Studies dealing with the influence of height on trunk strength have often focused on growing adolescents (Sinaki et al. 1996; Sunnegardh et al. 1988). We conclude that height is a strong factor influencing lifting capacity but the causality of this phenomenon is difficult to explain. Height has also been proposed to be a potential risk factor for LBP. Since there are conflicting results in this regard (Han et al. 1997; Leclerc et al. 2003; Nissinen et al. 1994), this relationship remains unclear. We conclude that height merits more attention in research on trunk function and the occurrence of LBP.

The maximum weight lifted in our study group was slightly larger for the PILE than for the FCE. This could be due partly to a learning effect since the PILE test was always performed the week after the FCE testing. Randomization of the lifting test sequence was not possible since the determination of limits by the therapist administering the FCE should not be affected by the capacities previously measured in the PILE test. There were significant differences between the lower and upper lifting tasks. The average maximum weight lifted in the two lower lifting tests was similar to the proposed weight limits for occasional lifting.

Our model, which included maximum force, endurance of the muscles involved in moment production and stabilization of the trunk during movements in the sagittal plane, cardiovascular endurance, mobility and coordination ability, predicted only 18–35% of the lifting capacity, depending on the test and the setting used. The fraction of variation in the lifting capacity that could be explained by our model was clearly smaller for the FCE than for the PILE. The small proportion of variance explained by our model might be due to the fact that our study group was rather homogeneous and no factors clearly constraining lifting capacity were present. Additionally, we hypothesize that lifting technique and motor control strongly contribute to lifting capacity. This is supported by the results of Dempsey et al. (1998) who showed a strong association of peak isoinertial power with maximum acceptable weight of lift (MAWL) whereas isometric and isokinetic strength, which do not make great demands on motor control, were not associated with the MAWL. Our parameter for coordination ability, based on postural sway and trunk rotation movements, did not correctly represent the aspects of motor control relevant for lifting tasks. Furthermore, we did not include strength measurements of the biceps brachii, which in our study group of women without severe back problems, could determine the lifting capacity of some subjects.

Finally, in this study we did not take into account any psychosocial factors. Jones and Kumar (2003) reported that social and work factors have a higher predictive value than physical factors. Basic functions of the gross motor system have low predictive power in regard to true work ability and to return to work (Hildebrandt et al. 1997). Furthermore, we showed that the basic functions tested did not predict the lifting capacity. The lifting tests were designed to mimic work situations and proved to predict work ability and probability for return to work (Matheson et al. 2002), but the lifting tasks performed during the FCEs should be similar to lifting tasks required on the job in order to improve the validity of testing (Hart et al. 1993). Therefore, these results give evidence to suggest the inclusion of the evaluation of lifting capacity in clinical practice. Furthermore, they raise questions about the predictive value of strength and endurance tests in regard to lifting capacity and work ability, and about the role of height in trunk function.

Acknowledgements The authors wish to thank Leanne Pobjoy for her help with the manuscript.

References

Astrand PO (1970) Textbook of work physiology. McGraw-Hill, New York

Biering-Sorensen F (1984) Physical measurements as risk indicators for low-back trouble over a one-year period. Spine 9:106–119

Brouwer B, Culham EG, Liston RA, Grant T (1998) Normal variability of postural measures: implications for the reliability of relative balance performance outcomes. Scand J Rehabil Med 30:131–137

Brouwer S, Reneman MF, Dijkstra PU, Groothoff JW, Schellekens JM, Goeken LN (2003) Test-retest reliability of the Isernhagen Work Systems Functional Capacity Evaluation in patients with chronic low back pain. J Occup Rehabil 13:207–218

Bunce SM, Hough AD, Moore AP (2004) Measurement of abdominal muscle thickness using M-mode ultrasound imaging during functional activities. Man Ther 9:41–44

Cholewicki J, Juluru K, McGill SM (1999) Intra-abdominal pressure mechanism for stabilizing the lumbar spine. J Biomech 32:13–17

Dempsey PG, Ayoub MM, Westfall PH (1998) Evaluation of the ability of power to predict low frequency lifting capacity. Ergonomics 41:1222–1241

Essendrop M, Schibye B, Hansen K (2001) Reliability of isometric muscle strength tests for the trunk, hands and shoulders. Int J Ind Erg 28:379–387

Essendrop M, Maul I, Laubli T, Riihimaki H, Schibye B (2002) Measures of low back function: a review of reproducibility studies. Clin Biomech 17:235–249

Gauvin MG, Riddle DL, Rothstein JM (1990) Reliability of clinical measurements of forward bending using the modified fingertip-to-floor method. Phys Ther 70:443–447

Gore CJ, Booth ML, Bauman A, Owen N (1999) Utility of pwc75% as an estimate of aerobic power in epidemiological and population-based studies. Med Sci Sports Exerc 31:348–351

Gorelick M, Brown JM, Groeller H (2003) Short-duration fatigue alters neuromuscular coordination of trunk musculature: implications for injury. Appl Ergon 34:317–325

Gross DP, Battie MC (2002) Reliability of safe maximum lifting determinations of a functional capacity evaluation. Phys Ther 82:364–371

Han TS, Schouten JS, Lean ME, Seidell JC (1997) The prevalence of low back pain and associations with body fatness, fat distribution and height. Int J Obes Relat Metab Disord 21:600–607

Hart DL, Isernhagen SJ, Matheson LN (1993) Guidelines for functional capacity evaluation of people with medical conditions. J Orthop Sports Phys Ther 18:682–686

Hildebrandt J, Pfingsten M, Saur P, Jansen J (1997) Prediction of success from a multidisciplinary treatment program for chronic low back pain. Spine 22:990–1001

Hyytiainen K, Salminen JJ, Suvitie T, Wickstrom G, Pentti J (1991) Reproducibility of nine tests to measure spinal mobility and trunk muscle strength. Scand J Rehabil Med 23:3–10

Isernhagen SJ (1991) Functional capacity evaluation and work hardening perspectives. In: Mayer TG, Mooney V, Gatchel R (eds) Contemporary conservative care for painful spinal disorders. Lea & Febinger, Philadelphia

Isernhagen SJ (1992) Functional capacity evaluation: rationale, procedure, utility of the kinesiophysical approach. J Occup Rehabil 2:157–168

Isernhagen SJ, Hart DL, Matheson LM (1999) Reliability of independent observer judgments of level of lift effort in a kinesiophysical functional capacity evaluation. Work 12:145–150

Ito T, Shirado O, Suzuki H, Takahashi M, Kaneda K, Strax TE (1996) Lumbar trunk muscle endurance testing: an inexpensive alternative to a machine for evaluation. Arch Phys Med Rehabil 77:75–79

Jones T, Kumar S (2003) Functional capacity evaluation of manual materials handlers: a review. Disabil Rehabil 25:179–191

Keller A, Hellesnes J, Brox JI (2001) Reliability of the isokinetic trunk extensor test, Biering-Sorensen test, and Astrand bicycle test: assessment of intraclass correlation coefficient and critical difference in patients with chronic low back pain and healthy individuals. Spine 26:771–777

Kippers V, Parker AW (1987) Toe-touch test: a measure of its validity. Phys Ther 67:1680–1684

Leclerc A, Tubach F, Landre MF, Ozguler A (2003) Personal and occupational predictors of sciatica in the GAZEL cohort. Occup Med (Lond) 53:384–391

Looze MP de, Groen H, Horemans H, Kingma I, van Dieen JH (1999) Abdominal muscles contribute in a minor way to peak spinal compression in lifting. J Biomech 32:655–662

Luk KD, Lu WW, Kwan WW, Hu Y, Wong YW, Law KK, Leong JC (2003) Isokinetic and isometric lifting capacity of Chinese in relation to the physical demand of the job. Appl Ergon 34:201–204

Matheson LN, Isernhagen SJ, Hart DL (2002) Relationships among lifting ability, grip force, and return to work. Phys Ther 82:249–256

Mayer TG, Barnes D, Kishino ND, Nichols G, Gatchel RJ, Mayer H, Mooney V (1988a) Progressive isoinertial lifting evaluation. I. A standardized protocol and normative database. Spine 13:993–997

Mayer TG, Barnes D, Nichols G, Kishino ND, Coval K, Piel B, Hoshino D, Gatchel RJ (1988b) Progressive isoinertial lifting evaluation. II. A comparison with isokinetic lifting in a disabled chronic low-back pain industrial population. Spine 13:998–1002

Mital A, Nicholson AS, Ayoub MM (1993) Manual materials handling. Taylor & Francis, London

Nissinen M, Heliovaara M, Seitsamo J, Alaranta H, Poussa M (1994) Anthropometric measurements and the incidence of low back pain in a cohort of pubertal children. Spine 19:1367–1370

Olivieri M, Denier-Bont F, Hallmark ML (1999) Evaluation der funktionellen Leistungsfähigkeit (EFL) nach Susan Isernhagen. Medizinische Mitteilungen 69:15–27

Sinaki M, Limburg PJ, Wollan PC, Rogers JW, Murtaugh PA (1996) Correlation of trunk muscle strength with age in children 5 to 18 years old. Mayo Clin Proc 71:1047–1054

Steinberg U, Windberg H (1994) Sonderschrift 9. Leitfaden Sicherheit und Gesundheitsschutz bei der manuellen Handhabung von Lasten. Bundesanstalt für Arbeitsmedizin, Bremerhaven

Suni JH, Oja P, Laukkanen RT, Miilunpalo SI, Pasanen ME, Vuori IM, Vartiainen TM, Bos K (1996) Health-related fitness test battery for adults: aspects of reliability. Arch Phys Med Rehabil 77:399–405

Sunnegardh J, Bratteby LE, Nordesjo LO, Nordgren B (1988) Isometric and isokinetic muscle strength, anthropometry and physical activity in 8- and 13-year-old Swedish children. Eur J Appl Physiol Occup Physiol 58:291–297

Troup JD, Foreman TK, Baxter CE, Brown D (1987) 1987 Volvo award in clinical sciences. The perception of back pain and the role of psychophysical tests of lifting capacity. Spine 12:645–657

Westhoff M (1994) The inter-observer reliability of a standardized protocol to measure the physical capacity of the musculoskeletal system. Faculty of Human Movement Sciences, Amsterdam Vrije Universiteit, Amsterdam